Acid-catalysed rearrangement of the sand fly pheromone sobralene to verticillenes, consolidating its relationship *inter alia* to the taxanes and phomactins

Matthew J. Palframan *[a]
Krishna K. Bandi [b]
James G. C.Hamilton [b]
Gerald Pattenden *[a]

[a] *School of Chemistry, The University of Nottingham, University Park, Nottingham, NG7 2RD, UK.*

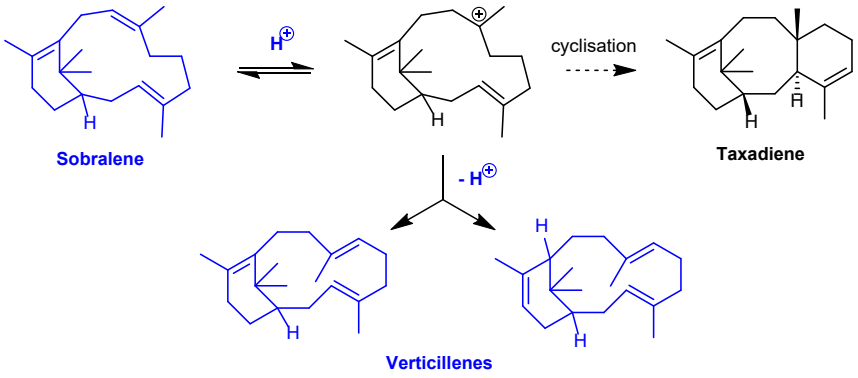
[b] *Division of Biomedical and Life Sciences, Faculty of Health and Medicine, Lancaster University, Lancaster, LA1 4YG, UK .*

* indicates the main/corresponding author.

Note: If an author has relocated from where the research was carried out, this is indicated with the next available reference number (usually 1); the new address should be added as that reference in the reference section.

e-mail gp@nottingham.ac.uk

Click here to insert a dedication.



Received:
Accepted:
Published online:
DOI:

Abstract The sex-aggregation pheromone sobralene **1**, produced by the sand fly *Lutzomyia longipalpis*, is isomerised to the verticillenes **2** and **5** in the presence of mild acid, thereby providing credence to the proposal that sobralene is a likely shunt metabolite of the taxadiene synthase cascade.

Key words pheromone, sand-fly, verticillenes, metabolite, taxadiene

Sobralene **1** is a novel macrocyclic diterpene hydrocarbon recently characterised as the sex-aggregation pheromone produced by populations of the sand fly *Lutzomyia longipalpis* from Sobral, Brazil.[1,2] The sand fly species complex *L. longipalpis* is the main carrier of the Protist parasite *Leishmania infantum*, the causative agent of American visceral leishmaniasis (AVL), which is a potentially fatal human disease in South and Central America.[3,4]

Sobralene **1** is the *cis* (*Z*-)-C8,C9 isomer of the *trans* (*E*)-C7,C8 diterpene hydrocarbon **2** known as "verticillene" which, for several decades, was proposed as a likely key intermediate in the biosynthesis of taxadiene **3**,[5] and later, of the phomactin family of platelet activating factor (PAF) antagonists , *e.g.* **4**.[6] However, although a number of alkene isomers of verticillene, *e.g.* **5**, *ent*-**5**, **6**, **7**, *ent*-**7**, have been characterised from various plant and other sources,[7,8,9,10] the particular 3,7,11-triene isomer **2** has not been found "free" in Nature.[11] In recent years however, in targeted engineering studies of taxadiene synthase (TXS) , Brück *et al* [12] demonstrated that particular mutants of TXS give rise to the production of verticillene **2** together with its isomers **5** and **6**.

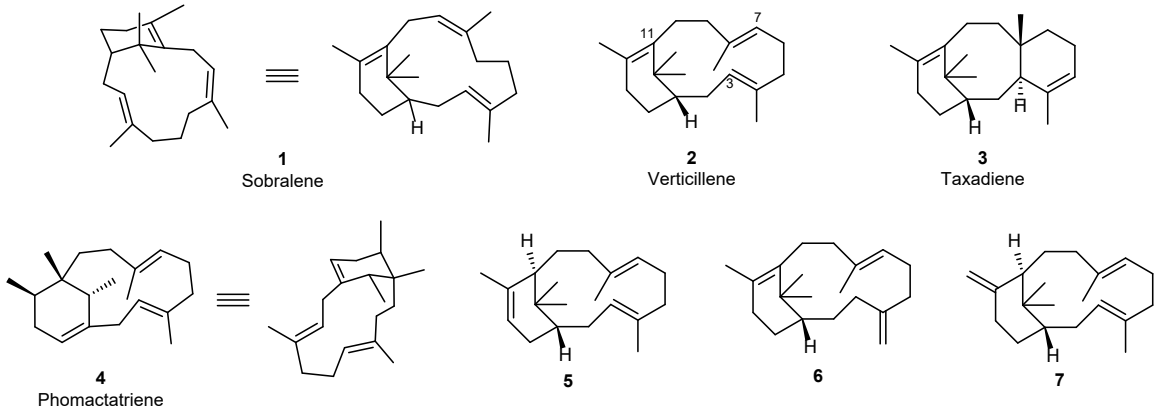
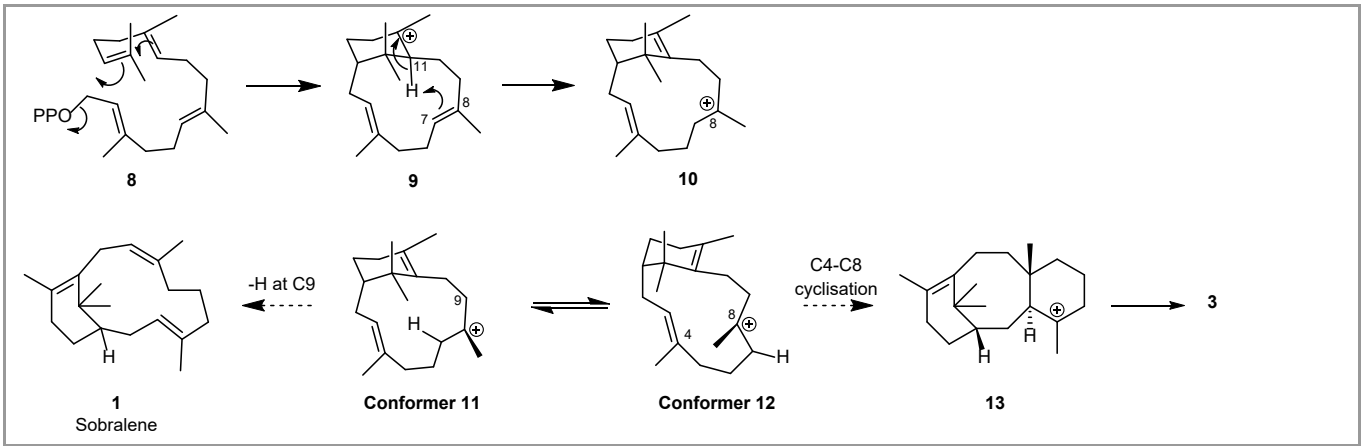


Figure 1 Sobralene **1**, and the structurally related verticillenes **2**, **5**, **6** and **7** , taxadiene **3** and phomactatriene **4**



Scheme 1 Proposal for the origin of sobralene **1** and taxadiene **3** from GGPP **8**, *via* the common carbocation intermediates **9** and **10**

It has been proposed that the 6,12-membered ring-fused system in sobralene **1** is derived in Nature from geranylgeranyl diphosphate (GGPP) **8** using a similar taxadiene synthase (TXS) cascade pathway to that leading to taxadiene **3**.[1] Thus, macrocyclisation of GGPP accompanied by transannulation first leads to the ring-fused carbocation intermediate **9** (Scheme 1). Extensive investigations have established that transfer of a proton from C11 to C7 in **9** leading to the corresponding C8 carbocation **10** is a key step in the biosynthesis of taxadiene. The carbocation **10** can assume a number of conformations, and quantum chemical calculations have reinforced an earlier proposal that a change in conformation of the carbocation **10** from **11** to **12**, is necessary prior to its transannulation to **13** *en route* to taxadiene **3**.[11] However, instead of transannulation, stereoselective elimination of the proton from C9, could lead to sobralene **1** (Scheme 1).[1]

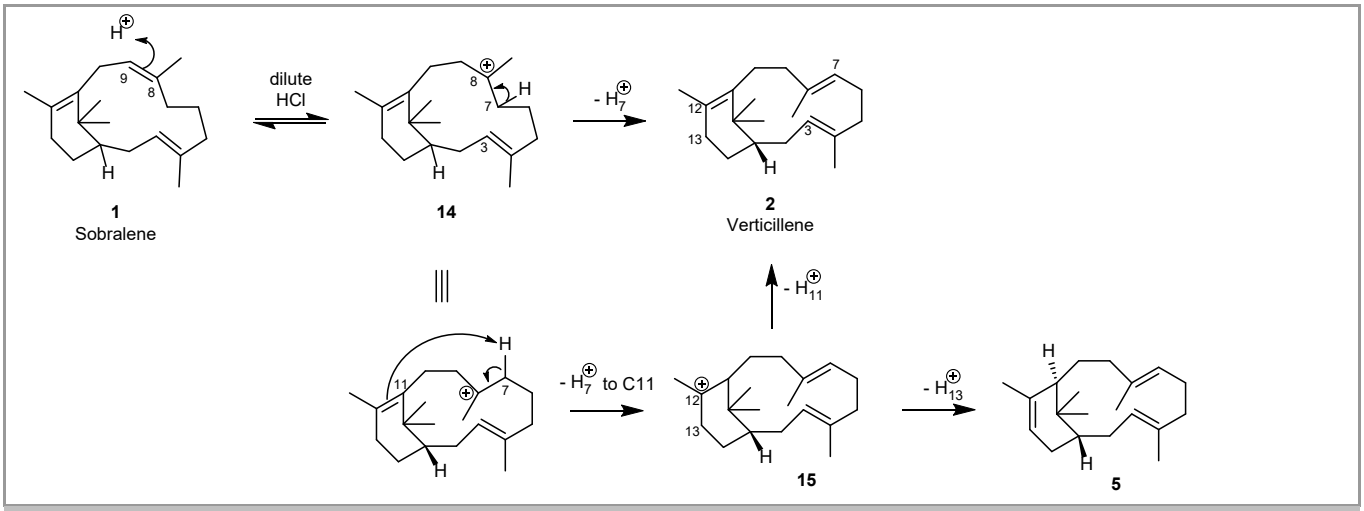
The interesting similarities in structure between the newly discovered sand fly pheromone sobralene **1**, and the verticillenes **2**, **5**-**7**, taxadiene **3** and phomactatriene **4,** prompted us to probe these relationships by undertaking a study of the rearrangement of sobralene under acid conditions. These studies are presented here.

Male sobral sand flies produce extremely minute quantities of pheromone secretions, and only at specific periods in their life cycle.[13,14] For our studies, we had batches of secretions which had been collected from many *thousand* sand flies over several months during different periods between 2014 and 2017, and then stored in different solvents. Fortunately, in a fresh sample of insect secretion (*ca.* 100 µg) collected in hexane from approx. two thousand male Sobral-2S sand flies, during four months in 2017, the pheromone was shown to constitute > 90% of the secretion when analysed by GC-MS (Figure 1 in ESI). Study of this sample by high field ^{1}H and ^{13}C NMR spectroscopy then permitted us to establish the structure **1** for the pheromone.[1]

We also had a larger sample of secretion (*ca.* 1 mg) collected from *ca.* 40,000 sand flies during 2014 but which had been stored in *deuteriochloroform* (*i.e.* mildly acidic conditions) for over 3 years. GC-MS analysis (Figure 2 in ESI) showed that this secretion contained five main constituents all of which had the same molecular ion at m/z 272, corresponding to $C_{20}H_{32}$. One of these compounds was present to the extent of *ca*. 36% (R_f 27.87) and a minor component (R_f 27.08, *ca.* 12%) corresponded to sobralene **1** from comparison of its GC-MS data with those of an authentic sample.

Examination of the EI mass spectrometry fragmentation data from the GC-MS analysis, alongside a library search on the NIST mass spectrometry database revealed that the fragmentation patterns for the main components in the insect secretion were all similar for isomers of the carbon backbone and substitution pattern observed for a "verticillene" hydrocarbon framework, *cf.* **5**. With this foresight we then examined the proton and carbon NMR spectra of the secretion. The major component in the secretion showed ^{1}H NMR data, *i.e.* δ 4.72 (dd, *J* 9.8, 1.0) and δ 5.19 (dd, *J* 12.0, 1.0) ppm which were diagnostic for the olefinic proton signals at C3 and at C7 respectively in the verticillene isomer **2** (Figure 3a in ESI). Likewise, other proton signals at δ 5.40 (brs), δ 5.28 (d, *J* 12. 0) and δ 4.82 (d, *J* 12. 0) in the secretion were superimposable on the C13, C3 and C7 olefinic proton resonances respectively in a synthetic sample of the isomeric verticillene **5** (Figure 3b in ESI). In addition, inspection and comparison of the data between δ 120 ppm and δ 140 ppm in the ^{13}C NMR spectrum of the secretion showed correlations with those found in authentic samples of the isomeric verticillenes **2** and **5** (Figures 4a and 4b in ESI). Thus, the sample showed olefinic carbon resonances corresponding to verticillene **2** at δ 136.0 (s), 132.8 (s), 131.5 (s), 128.5(d), 126.7(d) and 126.3 (s) ppm, and to the *endo* verticillene **5** at δ 136.0 (s), 133.1(s), 132.8 (s), 130.0(d), 124.9(d) and 121.8 (d) ppm (Figures 4a and 4b in ESI).

At this point in our studies we presumed that most of the sobralene **1** that had been present originally in the aforementioned decomposed secretion from the sand fly had undergone rearrangement *via* a sequence of proton exchanges leading to the isomeric verticillenes **2** and **5** during prolonged storage in the mildly acidic deuteriochloroform solution (c*f.* Scheme 2). To give credence to this supposition we carried out a controlled experiment of the isomerisation of a fresh sample of insect secretion containing >90% sobralene collected in hexane solvent, in the presence of mild acid. Thus, after transfer from hexane to deuteriochloroform a solution containing *ca*. 50 µg of sobralene (collected from *ca.* 1000 sand flies), was treated portion wise with a solution of 4N HCl in

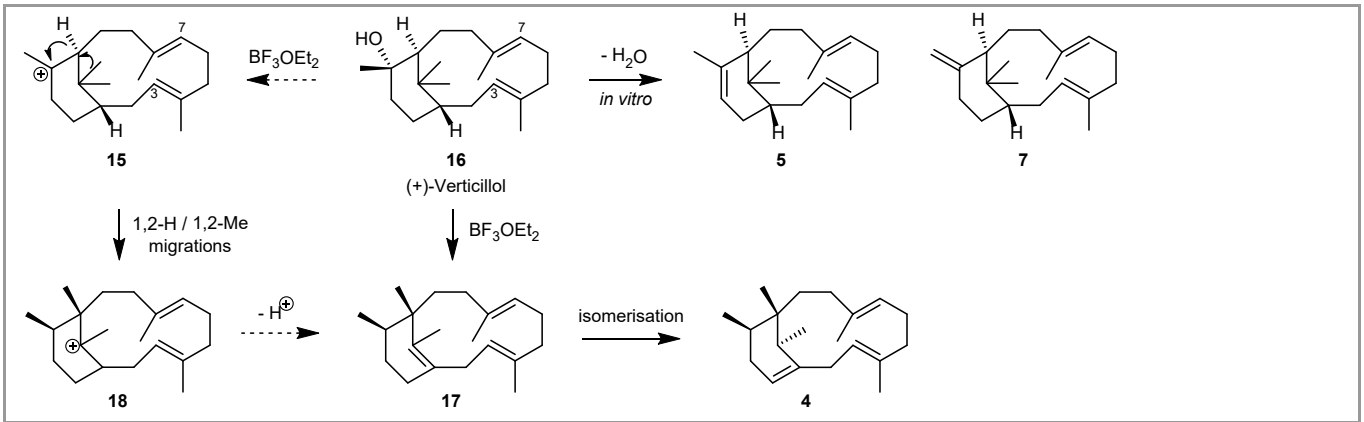


Scheme 2 Acid-catalysed rearrangement of sobralene **1** leading to the verticillene isomers **2** and **5**.

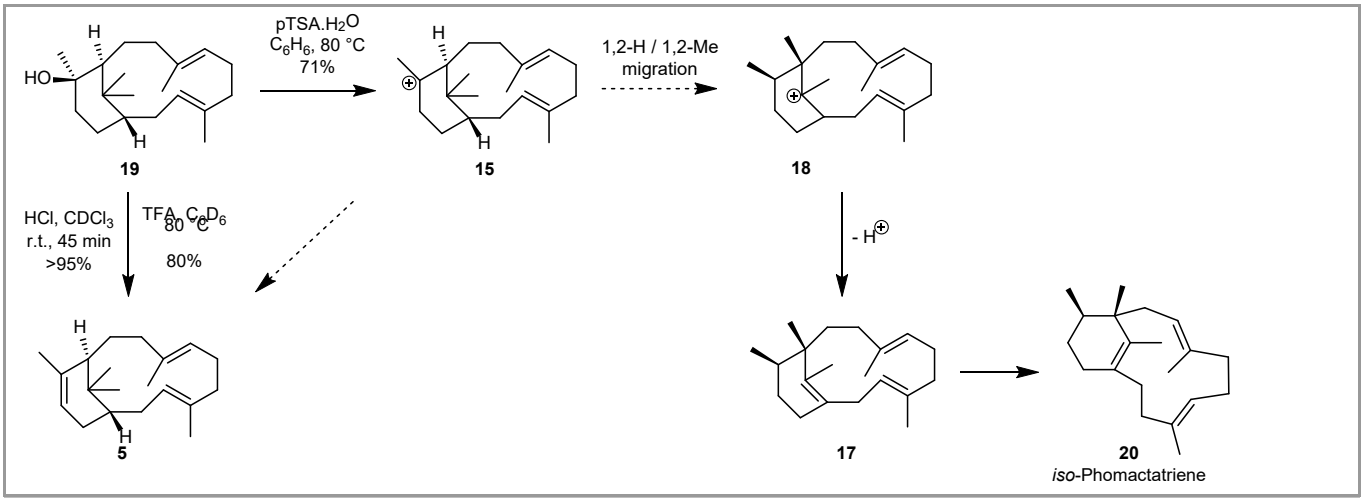
dioxane (1 µL through to 5µL) at room temperature, and the disappearance of sobralene was monitored by GC-MS analysis of samples taken at regular intervals (see Figures 5a, 5b, and 5c in ESI). When the amount of sobralene, as measured by peak area in GC-MS analysis, had diminished to approximately 15-20% of the total products having the same molecular ion m/z 272 ($C_{20}H_{32}$) in mass spectrometry analysis, the solution was neutralised by passing it through a plug of anhydrous K_2CO_3 located in a micropipette. The filtrate was then gently evaporated to dryness *in vacuo* and the residue was taken up in hexadeuterobenzene for further analysis by gas chromatography and NMR spectroscopy. GC-MS analysis showed the presence of one major product (>50%) and several minor products with m/z 272, in addition to a small amount of sobralene (< 15%) in the residue (Figure 5c in ESI). With the very small amount of material available, data between δ 6.0-10.0 and δ 0.5-3.50 ppm in the proton NMR spectrum of the product resulting from the acid treatment of sobralene were largely meaningless (Figure 6 in ESI). However, analysis of the corresponding olefinic proton region between δ 4.75- 5.5 ppm in the NMR spectrum showed that the major product in the sample was the verticillene isomer **2.** This structure followed from analysis and comparison of its olefinic proton NMR data with those for synthetic verticillene and comparisons of it MS data (Figure 7 in ESI).[15,16] Thus, in particular, the ^{1}H NMR spectrum (Figure 6 in ESI) exhibited resonance signals at δ 4.69 (dd, *J* 9.8, 1.0) and δ 5.15 (dd, *J* 12. 0, 1.0) ppm which are diagnostic for the olefinic proton signals at C3 and at C7 respectively in verticillene **2**. We later obtained a small sample of authentic verticillene **2** from colleagues in Munich, Germany [12] and demonstrated that its GC-MS data (Figure 7 in ESI) were identical with those we had recorded earlier for material resulting from treatment of sobralene **1** with 4N HCl in dioxane.

We rationalise the acid-catalysed isomerisation of sobralene **1** into the verticillene **2** occurring in a straightforward manner *via* protonation of the *cis* (*Z*)-C8,C9 alkene bond in **1** followed by loss of the α-orientated proton at C7 in the resulting C8 carbocation intermediate **14** (*cf*. **10**, Scheme 1) leading to the *trans* (*E*)- C7,C8 alkene bond in **2** (Scheme 2). We propose that the *endo* (C12-C13) verticillene isomer **5** identified in the earlier described acid treatment of pheromone secretion is derived from the same carbocation intermediate **14** following initial transfer of the proton at C7 to C11 leading to the C18 carbocation **15** (*cf.* 9 , Scheme 1), and then loss of a proton at C13 (Scheme 2). [17]

The alkene isomers **5** and **7** of verticillene **2** were first reported in 1964 as products resulting from the dehydration of the natural product (+)-verticillol **16**, isolated from the wood of *Sciadopitys vericillata* [15,16]. Later, in 2005, the same verticillenes **5** and **7** were isolated together with the 4- *exo*



Scheme 3 Rearrangement of verticillol **16** to phomactatriene **17** in the presence of BF_3 etherate



Scheme 4 Rearrangement of 12-*epi*-verticillol **19** to *iso*-phomactatriene **20** in the presence of pTSA.H_2O.

verticillene isomer **6** from plants of the genus *Bursera* endemic to Mexico.[9] At a similar time, during investigations of the treatment of verticillol **16** with boron trifluoride etherate at -78 °C Oikawa *et. al* [18] identified the interesting phomactatriene **17** in addition to the verticillenes **5** and **7** amongst the products**.** Indeed the same BF_3-catalysed reaction with verticillol at -30 °C led to **17** as the major product. In complementary studies Oikawa and his colleagues later identified the isomeric verticillenes **5** and **7** together with the phomactatriene **17** in mycelium extracts of the marine fungus *Phoma* sp. The phomactatriene **17** has its origins in the same C12 carbocation intermediate **15** proposed as an intermediate in the biosynthesis of both sobralene **1** and taxadiene **3** from GGPP **8** (*cf.* Scheme 1). Thus, a sequence of carbocation initiated 1,2 hydrogen and 1,2-methyl group shifts in **15** first leads to the rearranged cyclohexane carbocation **18** which, following elimination of a proton at C1, produces the phomactatriene **17** (Scheme 3).

As a corollary to our studies we undertook an independent study of the rearrangement of verticillol using different acid conditions to determine whether or not phomactatriene structures might be amongst the products of the rearrangement of sobralene **1** in acid. For these studies we were fortunate to obtain a sample of 12-*epi* verticillol **19** which had only recently been isolated from *Bursera microphylla* by Marcotullio *et al*.[10] Treatment of a solution of 12-*epi* verticillol **19** in $CDCl_3$ with 4N HCl in dioxan at room temperature, resulted in its complete conversion into *endo*-verticillene **5**, within 45 min, in essentially quantitative yield.[19] Similarly, treatment of **19** with TFA in benzene at 60 °C also led to the *endo*-verticillene in 80% yield, albeit more slowly (*i.e. ca* 5 h). By contrast however, treatment of a solution of 12-*epi* verticillol **19** with p-TSA.H_2O in benzene at 80 °C for 20 mins resulted in a facile rearrangement *via* the carbocation intermediates **15** and **18** leading to the formation of the interesting *iso*-phomactatriene **20** as the only isolable product in 71% yield.[20] *iso*-Phomactatriene **20** is a positional isomer of the known phomactatriene **17** where the alkene bonds at C2-4 and C7-8 in **17** have migrated to C3-4 and C8-9 respectively in **20** (Scheme 4). Interestingly, the phomactatriene isomer **20** was earlier reported amongst the products of treatment of (+)-verticillol **16,** isolated from *Sciadopits verticillate*, with lanthanum triflate , and the same authors also demonstrated that **17** could be isomerised to **20** in low yield following treatment with BF_3 etherate at -20 °C. Notwithstanding these interesting observations, however, we were not able to secure any evidence for the presence of phomactatriene-like structures amongst the products of acid-catalysed rearrangement of the pheromone sobralene **1** produced by the sand fly *Lutzomyia longipalpis*.

In summary we have demonstrated that the major products which result from treatment of the sex aggregation pheromone sobralene **1** of the sand fly *L. longipalpis* from Sobral, Brazil with dilute acid are the isomeric verticillenes **2** and **5**. The results reinforce our earlier suggestion of the close relationship between sobralene and verticillenes in Nature and, *inter alia*, the taxane **3** and phomactin **4** families of natural products. Interestingly, the processes which result in the formation of the verticillenes **2** and **5**, sobralene **1**, and taxadiene **3** (also phomactatriene **17**) in Nature share the common C8 verticillyl carbocation intermediate **10** (cf. **14**). This observation illustrates that very subtle changes in the conformation of **10** present several opportunities for its subsequent chemistry *in vivo*.

Acknowledgment

We thank the School of Chemistry at The University of Nottingham for support and for access to facilities. We also thank Professor Maria Carla Marcotullio (Dept. of Pharmaceutical Sciences, Univ. of Perugia , Italy) for a sample of 12-*epi*-verticillol **19** and Professor Monica Fuchs (Technical Univ. of Munich, Germany) for a sample of verticillene **2** for GC-MS comparison.

Supporting Information

YES (this text will be updated with links prior to publication)

Primary Data

NO (this text will be deleted prior to publication)

References and Notes

(1) M. J. Palframan, K. K. Bandi, J. G. C. Hamilton and G. Pattenden, *Tetrahedron Lett.,* 2018, **59**, 20, 1921.

(2) Two homosesquiterpene structures have previously been characterised as pheromones in other populations of sand fly from different regions of Brazil. See a) J.G.C. Hamilton, H.C. Ibbotson, A.M. Hooper, K. Mori, J.A. Pickett and S. Sano, *Chem. Commun.*, 1999, **8**, 2335; b) J.G.C. Hamilton, A.M. Hooper, K. Mori, J.A. Pickett and S. Sano, *Chem. Commun*., 1999, **4**, 355.

(3) L. Deane and M. Deane, *Rev. Inst. Med. Trop*., 1962, **4**, 198.
(4) R. Lainson and E.F. Rangel, *Mem. Inst. Oswaldo Cruz*, 2005, **100**, 811.
(5) J. W. Harrison, R. M. Scrowston and B. Lythgoe, *J. Chem. Soc. C*, 1966, 1933–1945.
(6) For recent reviews see : *a*) J. Ciesielskia, and A. Frontiera, *Organic Preparations and Procedures International: The New Journal for Organic Synthesis*, 2014, **46**, 214-251; b) W. P. D. Goldring and G. Pattenden, *Acc. Chem. Res.*, 2006, **39**, 354 – 361.
(7) Y. Asakawa, *Progress in the Chemistry of Organic Natural Products* , Eds. W. Herz, G. W. Kirby, R. E. Moore, W. Steglich and C. Tamm, Springer-Verlag: Vienna, 1995, **65**, 279–280.
(8) S. Basar, A. Koch, and W. A. König, *Flavour Fragr. J.*, 2001, **16**, 315–318.
(9) J. D. Hernández-Hernández, L. U. Román-Marín, C. M.; Cerda-García-Rojas and P. Joseph-Nathan, *J. Nat. Prod.*, 2005, **68**, 1598–1602.
(10) F. Messina, M. Curini, C. D. Sano, C. Zadra, G. Gigliarelli, L. A. Rascón-Valenzuela, R. E. R. Zepeda and M. C. Marcotullio, *J. Nat. Prod.*, 2015, **78**, 1184-1188.
(11) For a recent perspective on the role of verticillenes in the biosynthesis of the taxanes and the phomactins see: M. J. Palframan and G. Pattenden, *Nat. Prod. Rep.*, 2019, **36**, 108-121.
(12) a) P. Schrepfer, A. Buettner, C. Goerner, M. Hertel, J. van Rijn, Frank Wallrapp, W Eisenreich, V. Sieber, Robert Kourist and T. Brück, *PNAS*, 2016, E958; b) K. Kemper, M. Hirte, M. Reinbold, M. Fuchs and T. Brück, *Beilstein J. Org. Chem.*, 2017, **13**, 845 - 854, see also : S. Edgar, F.-S. Li, K. Qiao, J. K. Weng and G. Stephanopoulos, *ACS Synt. Biol.*, 2017, **6**, 201-205.
(13) R. Lane, A. Phillips, D. H. Molyneux, G. Procter, R. D. Ward. *Ann Trop Med Parasitol.,* 1985, **79**, 225–229.
(14) J. G. C. Hamilton, R. P. Brazil, R. Maingon, *J. Med. Entomol.*, 2004, **41**, 1021–1026.
(15) H. Erdtman, T. Norin, M. Sumimoto and A. Morrison, *Tetrahedron Lett*., 1964, **5**, 3879–3886 and B. Karlsson, A. M. Pilotti, A. C. Söderholm, T. Norin, and S. Sundin, *Tetrahedron,* 1978, **34**, 2349-2354.
(16) a) C. B. Jackson, and G. Pattenden, *Tetrahedron Lett.* 1985, **26**, 3393–3396; b) M. J. Begley, C. B. Jackson and G. Pattenden, *Tetrahedron*, 1990, **46**, 4907–4924; *cf.* T. Kato, M. Hoshikawa, Y. Yoshihiro, K. Izumi, Y. Uotsu, and K. Sakai, *Tetrahedron*, 2002, **58**, 9213, and Q. Jin, D. C. Williams, M. Hezari, R. Croteau and R. M. Coates, *J. Org. Chem.*, 2005, **70**, 4667–4675.
(17) Further studies are underway to elucidate whether a one-step transfer of the proton at C7 to C11 occurs or if two discrete steps are involved.
(18) T. Tokiwano, T. Endo, T. Tsukagoshi, H. Goto, E. Fukushic, and H. Oikawa, *Org. Biomol. Chem.*, 2005, **3**, 2713 – 2722.
(19) Trifluoroacetic acid (1.0μL, 0.0130 mmol) was added to a solution of 12-*epi* verticillol **19** (1.20 mg, 0.00413 mmol) in C_6D_6 (0.75 mL) and the resulting solution was heated at 60 °C for 3 hours. A further portion of trifluoroacetic acid (2.0 uL, 0.0260 mmol) was added, and the solution was heated for a further 2 hours. The solution was cooled to room temperature and evaporated to dryness. The residue was purified by silica gel chromatography, eluting with pentane to afford *endo*-verticillene (0.90 mg, 80%) as a pale yellow oil; ^{1}H-NMR (500 MHz, C_6D_6) 5.48 (1 H, bs, H13), 5.45 (1 H, d, *J* = 12.4, H3), 4.89 (1 H, d, *J* = 10.5, H7), 3.08 (1 H, bs, H11), 2.65 (1 H, ddd, *J* = 15.9, 11.9, 4.4, H2a), 2.49 (1 H, bm, H14a), 2.39 (1 H, m, H6a), 2.24 (1 H, dd, *J* = 12.6, 6.1, H9a), 2.10 (2 H, m, H9b and H5a), 1.97-1.90 (2 H, m, H5b and H6b), 1.90-1.81 (2 H, m, H2b and H14b),1.83 (3 H, s, H18), 1.52 (3 H, s, H20), 1.45 (3 H, s, H19), 1.42 (1 H, m, H1), 1.33-1.29 (2 H, m, H10), 0.88 (3 H, s, H16), 0.81 (3 H, s, H17); ^{13}C NMR (126 MHz, C_6D_6) δ 136.9 (C12), 132.8 (C8), 132.7 (C4), 130.4 (C7), 125.1 (C3), 122.2 (C13), 42.9 (C1), 41.3 (C5), 40.0 (C9), 38.4 (C11), 36.0 C15), 34.5 (C2), 31.2 (C14), 27.4 (C17), 27.1 (C6), 24.0 (C16), 23.4 (C18), 21.9 (C10), 15.8 (C19), 15.3 (C20). GCMS ($C_{20}H_{32}^+$, M^+) requires 272.2499; HR-EI-*m/z* found 272.2511.
(20) pTSA.H_2O (1.5mg, 0.00517 mmol) was added to a solution of 12-*epi* verticillol **19** (1.5 mg, 0.00517 mmol) in C_6D_6 (0.75 mL), and the resulting solution was heated at 80 °C for 20 minutes. The solution was cooled to room temperature, then passed through a short plug of silica, eluting with Et_2O (1 mL) and evaporated. The residue was purified by silica gel chromatography, eluting with pentane to give the *iso*-phomactatriene (1.0 mg, 71%) as a pale yellow oil; ^{1}H-NMR (500 MHz, C_6D_6) 5.02 (1H, m, H9), 4.77 (1H, d, *J* = 10.3 Hz, H5), 2.60 (1H, dt, *J* = 2.4, 13.2 Hz, H2a), 2.28 (1H, m, H6a), 2.15–2.05 (3H, m, H3a, H7a, H14a), 2.03 (2H, d, *J* = 8.1 Hz, H10), 2.00–1.98 (3H, m, H3b, H6b, H7b), 1.68–1.58 (3H, m, H2b, H12, H14b), 1.57 (3H, s, C4CH_3), 1.53 (3H, s, C4CH_3), 1.42 (3H, s, C15CH_3), 1.38 (2H, m, H13), 0.92 (3H, d, J = 6.9 Hz, C12CH_3), 0.90 (3H, s, C11CH_3); ^{13}C-NMR (125 MHz, C_6D_6) 133.3 (C15), 132.9 (C8), 132.1 (C4), 131.2 (C1), 127.8 (C5), 123.5 (C9), 43.2 (C11), 39.0 (C7), 37.9 (C3), 34.6 (C12), 34.4 (C10), 32.9 (C2), 31.3(C14), 28.0 (C13), 25.7 (C6), 20.0 (C11$\underline{C}H_3$), 17.3 (C12$\underline{C}H_3$), 16.9 (C8$\underline{C}H_3$), 16.4 (C4$\underline{C}$H3), 14.5 (C15$\underline{C}H_3$); GCMS ($C_{20}H_{32}^+$, M^+) requires 272.2499; HR-EI-*m/z* found 272.2508